



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/731,875  
Source: OIPC  
Date Processed by STIC: 12-30-03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efb/efs/downloads/documents.htm>> , EFS Submission User Manual - cPAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

## Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 10/731,875
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 _____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 _____ Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 _____ Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 _____ Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 _____ Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 _____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 _____ Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 _____ Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 _____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 <input checked="" type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 _____ Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 00/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 _____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 _____ Misuse of n/Xaa	"n" can only represent a single <u>nucleotide</u> ; "Xaa" can only represent a single <u>amino acid</u>	



IFWO

## RAW SEQUENCE LISTING

DATE: 12/30/2003

PATENT APPLICATION: US/10/731,875

TIME: 15:35:37

Input Set : A:\Thomas, et al. (9692-000029) Seq. Listing.ST25.txt

Output Set: N:\CRF4\12302003\J731875.raw

3 <110> APPLICANT: Applera Corporation  
 4 Thomas, Paul  
 5 Kejariwal, Anish  
 6 Campbell, Michael  
 7 Mi, Huaiyu  
 8 Diemer, Karen  
 9 Guo, Nan  
 10 Ladunga, Istvan  
 11 Lazareva, Betty  
 12 Muruganujan, Anushya  
 13 Rabkin, Steven  
 14 Vandergriff, Jody  
 15 Doremieux, Olivier  
 17 <120> TITLE OF INVENTION: Browsable Database For Biological Use  
 19 <130> FILE REFERENCE: 9692-000029  
 C--> 21 <140> CURRENT APPLICATION NUMBER: US/10/731,875  
 C--> 21 <141> CURRENT FILING DATE: 2003-12-09

**Does Not Comply  
Corrected Diskette Needed**

21 <160> NUMBER OF SEQ ID NOS: 63  
 23 <170> SOFTWARE: PatentIn version 3.2  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 62  
 27 <212> TYPE: PRT  
 28 <213> ORGANISM: Exemplary Protein Sequence  
 30 <400> SEQUENCE: 1  
 32 Leu Ala Val Glu Val Leu Phe Ile Leu Asp Ile Val Leu Asn Phe Arg  
 33 1 5 10 15  
 36 Thr Thr Phe Val Ser Lys Gly Gln Val Val Phe Ala Pro Lys Ile Cys  
 37 20 25 30  
 40 Leu His Tyr Val Thr Thr Trp Phe Leu Leu Asp Val Ile Ala Ala Leu  
 41 35 40 45  
 44 Pro Phe Asp Leu Leu His Ala Lys Val Asn Tyr Phe Gly Ala  
 45 50 55 60

P.6

See item 10 on error summary sheet.

48 <210> SEQ ID NO: 2  
 49 <211> LENGTH: 62  
 50 <212> TYPE: PRT  
 51 <213> ORGANISM: Exemplary Protein Sequence  
 53 <400> SEQUENCE: 2  
 55 Leu Ala Val Glu Val Leu Phe Ile Leu Asp Ile Val Leu Asn Phe Arg  
 56 1 5 10 15  
 59 Thr Thr Phe Val Ser Lys Gly Gln Val Val Phe Ala Pro Lys Ile Cys  
 60 20 25 30  
 63 Leu His Tyr Val Thr Thr Trp Phe Leu Leu Asp Val Ile Ala Ala Leu  
 64 35 40 45

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/731,875

DATE: 12/30/2003

TIME: 15:35:37

Input Set : A:\Thomas, et al. (9692-000029) Seq. Listing.ST25.txt

Output Set: N:\CRF4\12302003\J731875.raw

```

67 Pro Phe Asp Leu Leu His Ala Lys Val Asn Tyr Val Gly Ala
68      50                      55                      60
71 <210> SEQ ID NO: 3
72 <211> LENGTH: 62
73 <212> TYPE: PRT
74 <213> ORGANISM: Exemplary Protein Sequence
76 <400> SEQUENCE: 3
78 Leu Ala Val Glu Val Leu Phe Ile Leu Asp Ile Val Leu Asn Phe Arg
79 1      5                      10                      15
82 Thr Thr Phe Val Ser Lys Gly Gln Val Val Phe Ala Pro Lys Ile Cys
83      20                      25                      30
86 Leu His Tyr Val Thr Thr Trp Phe Leu Leu Asp Val Ile Ala Ala Leu
87      35                      40                      45
90 Pro Phe Asp Leu Leu His Ala Lys Val Asn Tyr Val Gly Ala
91      50                      55                      60
94 <210> SEQ ID NO: 4
95 <211> LENGTH: 62
96 <212> TYPE: PRT
97 <213> ORGANISM: Exemplary Protein Sequence
99 <400> SEQUENCE: 4
101 Ile Ala Val Glu Ile Leu Phe Ile Ile Asp Ile Ile Leu Asn Phe Arg
102 1      5                      10                      15
105 Thr Thr Tyr Val Ser Lys Gly Gln Val Ile Phe Glu Ala Arg Ile Cys
106      20                      25                      30
109 Ile His Tyr Val Thr Thr Trp Phe Ile Ile Asp Leu Ile Ala Ala Leu
110      35                      40                      45
113 Pro Phe Asp Leu Leu Tyr Ala Asn Val Thr Val Ser Leu Val
114      50                      55                      60
117 <210> SEQ ID NO: 5
118 <211> LENGTH: 62
119 <212> TYPE: PRT
120 <213> ORGANISM: Exemplary Protein Sequence
122 <400> SEQUENCE: 5
124 Ile Ala Val Glu Ile Leu Phe Ile Ile Asp Ile Ile Leu Asn Phe Arg
125 1      5                      10                      15
128 Thr Thr Tyr Val Ser Lys Gly Gln Val Ile Phe Glu Ala Arg Ile Cys
129      20                      25                      30
132 Thr His Tyr Val Thr Thr Trp Phe Ile Ile Asp Leu Ile Ala Ala Leu
133      35                      40                      45
136 Pro Phe Asp Leu Leu Tyr Ala Asn Val Thr Val Ser Leu Val
137      50                      55                      60
140 <210> SEQ ID NO: 6
141 <211> LENGTH: 62
142 <212> TYPE: PRT
143 <213> ORGANISM: Exemplary Protein Sequence
145 <400> SEQUENCE: 6
147 Ile Ala Val Glu Met Leu Phe Ile Leu Asp Ile Ile Leu Asn Phe Arg
148 1      5                      10                      15
151 Thr Thr Tyr Val Ser Gln Gly Gln Val Val Ser Ala Pro Arg Ile Gly

```

Same error

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/731,875

DATE: 12/30/2003

TIME: 15:35:37

Input Set : A:\Thomas, et al. (9692-000029) Seq. Listing.ST25.txt

Output Set: N:\CRF4\12302003\J731875.raw

```

152          20          25          30
155 Leu His Tyr Leu Ala Thr Trp Phe Phe Val Asp Leu Ile Ala Ala Leu
156          35          40          45
159 Pro Phe Asp Leu Leu Tyr Val Asn Ile Thr Thr Ser Leu Val
160          50          55          60
163 <210> SEQ ID NO: 7
164 <211> LENGTH: 62
165 <212> TYPE: PRT
166 <213> ORGANISM: Exemplary Protein Sequence
168 <400> SEQUENCE: 7
170 Ile Ala Val Glu Met Leu Phe Ile Leu Asp Ile Ile Leu Asn Phe Arg
171 1          5          10          15
174 Thr Thr Tyr Val Ser Gln Gly Gln Val Val Ser Ala Pro Arg Ile Gly
175          20          25          30
178 Leu His Tyr Leu Ala Thr Trp Phe Phe Val Asp Leu Ile Ala Ala Leu
179          35          40          45
182 Pro Phe Asp Leu Leu Tyr Val Asn Ile Thr Thr Ser Leu Val
183          50          55          60
186 <210> SEQ ID NO: 8
187 <211> LENGTH: 62
188 <212> TYPE: PRT
189 <213> ORGANISM: Exemplary Protein Sequence
191 <400> SEQUENCE: 8
193 Ile Ala Val Glu Met Leu Phe Ile Leu Asp Ile Ile Leu Asn Phe Arg
194 1          5          10          15
197 Thr Thr Tyr Val Ser Gln Gly Gln Val Ile Ser Ala Pro Arg Ile Gly
198          20          25          30
201 Leu His Tyr Leu Ala Thr Trp Phe Phe Ile Asp Leu Ile Ala Ala Leu
202          35          40          45
205 Pro Phe Asp Leu Leu Tyr Ile Asn Ile Thr Thr Ser Leu Val
206          50          55          60
209 <210> SEQ ID NO: 9
210 <211> LENGTH: 62
211 <212> TYPE: PRT
212 <213> ORGANISM: Exemplary Protein Sequence
214 <400> SEQUENCE: 9
216 Ile Ala Val Glu Met Leu Phe Ile Leu Asp Ile Ile Leu Asn Phe Arg
217 1          5          10          15
220 Thr Thr Tyr Val Ser Gln Gly Gln Val Ile Ser Ala Pro Arg Ile Gly
221          20          25          30
224 Leu His Tyr Leu Ala Thr Trp Phe Phe Ile Asp Leu Ile Ala Ala Leu
225          35          40          45
228 Pro Phe Asp Leu Leu Tyr Ile Asn Ile Thr Thr Ser Leu Val
229          50          55          60
232 <210> SEQ ID NO: 10
233 <211> LENGTH: 67
234 <212> TYPE: PRT
235 <213> ORGANISM: Exemplary Protein Sequence
237 <400> SEQUENCE: 10

```

same error,  
see item 10 on error  
summary report.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/731,875

DATE: 12/30/2003

TIME: 15:35:37

Input Set : A:\Thomas, et al. (9692-000029) Seq. Listing.ST25.txt

Output Set: N:\CRF4\12302003\J731875.raw

```

239 Ser Ile Val Asp Val Ile Phe Phe Ile Asp Ile Val Leu Asn Phe His
240 1          5          10          15
243 Thr Thr Phe Val Gly Pro Gly Gly Glu Val Val Ser Asp Pro Lys Val
244          20          25          30
247 Ile Met Asn Tyr Leu Lys Ser Trp Phe Ile Ile Asp Leu Leu Ser Cys
248          35          40          45
251 Leu Pro Asp Val Phe Asn Ala Phe Asp Arg Asp Glu Asp Gly Ile Gly
252          50          55          60
255 Ser Leu Phe

```

256 65

259 &lt;210&gt; SEQ ID NO: 11

260 &lt;211&gt; LENGTH: 67

261 &lt;212&gt; TYPE: PRT

262 &lt;213&gt; ORGANISM: Exemplary Protein Sequence

264 &lt;400&gt; SEQUENCE: 11

```

266 Ser Ile Val Asp Val Ile Phe Phe Ile Asp Ile Val Leu Asn Phe His
267 1          5          10          15
270 Thr Thr Phe Val Gly Pro Gly Gly Glu Val Val Ser Asp Pro Lys Val
271          20          25          30
274 Ile Met Asn Tyr Leu Lys Ser Trp Phe Ile Ile Asp Leu Leu Ser Cys
275          35          40          45
278 Leu Pro Asp Val Phe Asn Ala Phe Asp Arg Asp Glu Asp Gly Ile Gly
279          50          55          60

```

282 Ser Leu Phe

283 65

286 &lt;210&gt; SEQ ID NO: 12

287 &lt;211&gt; LENGTH: 67

288 &lt;212&gt; TYPE: PRT

289 &lt;213&gt; ORGANISM: Exemplary Protein Sequence

same error

291 &lt;400&gt; SEQUENCE: 12

```

293 Ser Ile Val Asp Val Ile Phe Phe Ala Asp Ile Leu Leu Asn Phe His
294 1          5          10          15
297 Thr Thr Phe Val Gly Pro Gly Gly Glu Val Val Ile Glu Pro Ser Val
298          20          25          30
301 Ile Gln Asn Tyr Phe Lys Ser Trp Phe Leu Ile Asp Leu Leu Ser Cys
302          35          40          45
305 Leu Pro Asp Ile Phe Tyr Met Phe Lys Arg Asp Asp Glu Arg Ile Gly
306          50          55          60

```

309 Ser Leu Phe

310 65

313 &lt;210&gt; SEQ ID NO: 13

314 &lt;211&gt; LENGTH: 67

315 &lt;212&gt; TYPE: PRT

316 &lt;213&gt; ORGANISM: Exemplary Protein Sequence

318 &lt;400&gt; SEQUENCE: 13

```

320 Ser Ile Val Asp Val Ile Phe Phe Ala Asp Ile Leu Leu Asn Phe His
321 1          5          10          15
324 Thr Thr Phe Val Gly Pro Gly Gly Glu Val Val Ile Glu Pro Ser Val
325          20          25          30

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/731,875

DATE: 12/30/2003

TIME: 15:35:37

Input Set : A:\Thomas, et al. (9692-000029) Seq. Listing.ST25.txt

Output Set: N:\CRF4\12302003\J731875.raw

```

328 Ile Gln Asn Tyr Phe Lys Ser Trp Phe Leu Ile Asp Leu Leu Ser Cys
329      35                      40                      45
332 Leu Pro Asp Ile Phe Tyr Met Phe Lys Arg Asp Asp Glu Arg Ile Gly
333      50                      55                      60
336 Ser Leu Phe
337 65
340 <210> SEQ ID NO: 14
341 <211> LENGTH: 67
342 <212> TYPE: PRT
343 <213> ORGANISM: Exemplary Protein Sequence
345 <400> SEQUENCE: 14
347 Ser Val Val Asp Val Ile Phe Leu Val Asp Ile Val Leu Asn Phe His
348 1      5                      10                      15
351 Thr Thr Phe Val Gly Pro Gly Gly Glu Val Ile Ser Asp Pro Lys Leu
352      20                      25                      30
355 Ile Met Asn Tyr Leu Lys Thr Trp Phe Val Ile Asp Leu Leu Ser Cys
356      35                      40                      45
359 Leu Pro Asp Ile Ile Asn Ala Phe Glu Asn Val Asp Glu Gly Ile Ser
360      50                      55                      60
363 Ser Leu Phe
364 65
367 <210> SEQ ID NO: 15
368 <211> LENGTH: 67
369 <212> TYPE: PRT
370 <213> ORGANISM: Exemplary Protein Sequence
372 <400> SEQUENCE: 15
374 Ser Val Val Asp Val Ile Phe Leu Val Asp Ile Val Leu Asn Phe His
375 1      5                      10                      15
378 Thr Thr Phe Val Gly Pro Gly Gly Glu Val Ile Ser Asp Pro Lys Leu
379      20                      25                      30
382 Ile Met Asn Tyr Leu Lys Thr Trp Phe Val Ile Asp Leu Leu Ser Cys
383      35                      40                      45
386 Leu Pro Asp Ile Ile Asn Ala Phe Glu Asn Val Asp Glu Gly Ile Ser
387      50                      55                      60
390 Ser Leu Phe
391 65
394 <210> SEQ ID NO: 16
395 <211> LENGTH: 80
396 <212> TYPE: PRT
397 <213> ORGANISM: Exemplary Protein Sequence
399 <400> SEQUENCE: 16
401 Ser Ile Val Asp Val Ile Phe Leu Val Asp Ile Val Leu Asn Phe His
402 1      5                      10                      15
405 Thr Thr Phe Val Gly Pro Ala Gly Glu Val Ile Ser Asp Pro Lys Leu
406      20                      25                      30
409 Ile Met Asn Tyr Leu Lys Thr Trp Phe Val Ile Asp Leu Leu Ser Cys
410      35                      40                      45
413 Leu Pro Asp Val Ile Asn Ala Phe Glu Asn Val Asp Glu Val Ser Ala
414      50                      55                      60

```

The type of errors shown exist throughout  
the Sequence Listing. Please check subsequent  
sequences for similar errors.

## RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 12/30/2003

PATENT APPLICATION: US/10/731,875

TIME: 15:35:38

Input Set : A:\Thomas, et al. (9692-000029) Seq. Listing.ST25.txt

Output Set: N:\CRF4\12302003\J731875.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:50; Xaa Pos. 8

Seq#:60; Xaa Pos. 5



## VERIFICATION SUMMARY

DATE: 12/30/2003

PATENT APPLICATION: US/10/731,875

TIME: 15:35:38

Input Set : A:\Thomas, et al. (9692-000029) Seq. Listing.ST25.txt

Output Set: N:\CRF4\12302003\J731875.raw

L:21 M:270 C: Current Application Number differs, Replaced Current Application No

L:21 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:1481 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0

L:1597 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60 after pos.:0